# Title: Variability of spatial aggregation effects of climate data on regional yield simulation by crop models for a selected region in Germany

**Suggested running page head: Crop model climate input data aggregation effects**

*Hoffmann, Holger[1]; Zhao, Gang[1*]; van Bussel, Lenny G.J.[1,2]; Enders, Andreas[1]; Specka, Xenia[3]; Sosa, Carmen[4]; Yeluripati, Jagadeesh[5, 17]; Tao, Fulu[6]; Constantin, Julie[7]; Raynal, Helene[7]; Teixeira, Edmar[8]; Grosz, Balázs[9]; Doro, Luca[10]; Zhao, Zhigan[11]; Wang, Enli[11]; Nendel, Claas[3]; Kersebaum, Kurt-Christian[3]; Haas, Edwin[12]; Kiese, Ralf[12]; Klatt, Steffen[12]; Eckersten, Henrik[13]; Vanuytrecht, Eline[14]; Kuhnert, Matthias[5]; Lewan, Elisabet[4]; Rötter, Reimund[6]; Roggero, Pier Paolo[10]; Wallach, Daniel[7]; Cammarano, Davide[16], Asseng, Senthold[16]; Krauss, Gunther[1]; Siebert, Stefan[1]; Gaiser, Thomas[1]; Ewert, Frank[1]*

[1] Crop Science Group, Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Katzenburgweg 5, 53115 Bonn, DE

[2] Plant Production Systems Group, Wageningen University, P.O. Box 430, 6700 AK, Wageningen, NL

[3] Institute of Landscape Systems Analysis, Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, DE

[4] Biogeophysics and water quality, Department of Soil and Environment, Swedish University of Agricultural Sciences, Lennart Hjelms väg 9, 750 07 Uppsala, SE

[5] Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, 23 St Machar Drive, Aberdeen AB24 3 UU, Scotland, UK

[6] Climate Impacts Group, Natural Resources Institute Finland (Luke), 00790 Helsinki, FI

[7] INRA, UMR 1248 AGIR & UR0875 MIA-T, F-31326 Auzeville, FR

[8] Systems Modelling Team (Sustainable Production Group), The New Zealand Institute for Plant and Food Research Limited, Canterbury Agriculture & Science Centre, Gerald St, Lincoln 7608, NZ

[9] Thünen-Institute of Climate-Smart-Agriculture, Bundesallee 50, 38116 Braunschweig, DE

[10] Desertification Research Group, Universitá degli Studi di Sassari, Viale Italia 39, 07100 Sassari, IT

[11] CSIRO Land and Water, Clunies Ross Street, Canberra, ACT, AU

[12] Institute of Meteorology and Climate Research – Atmospheric Environmental Research, Karlsruhe Institute of Technology, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, DE

[13] Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 750 07 Uppsala, SE

[14] Division Soil & Water Management, KU Leuven, Celestijnenlaan 200E, PO 2411, 3001 Heverlee, BE

[16] Agricultural & Biological Engineering Department, University of Florida, Frazier Rogers Hall, Gainesville, FL 32611, USA

[17] The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

*Corresponding author: Holger Hoffmann, Tel: (+49) 228 73 2876, fax: (+49) 228 73 2870, email: hhoffmann@uni-bonn.de

**Abstract** Field-scale crop models are often applied at coarser spatial resolutions than the field. However, little is known on the response of the models to spatially aggregated climate input data and why these responses can differ across models. Depending on the model, regional yield estimates from large-scale simulations may therefore be biased as compared to simulations with high-resolution input data.

We therefore evaluated this so-called aggregation effect for 13 crop models for a selected region in Germany. For this purpose, the models were supplied with climate data of 1 km resolution and spatial aggregates of up to 100 km resolution raster. The models were used with two crops (winter wheat and silage maize) and three production situations (potential, water limited and nitrogen-water-limited growth) to improve the understanding of errors in model simulations related to data aggregation and possible interactions with the model structure. The most important climate variables identified to determining the model-specific input data aggregation on simulated yields were mainly related to changes in radiation (winter wheat) and temperature (silage maize). Additionally, aggregation effects were systematic since models differed in the systematic fraction of the aggregation effect, regardless of the extent of the effect (20 to 66 % as compared to 1.7 % for random effects). Climate input data aggregation changed the mean simulated regional yield up to 0.2 t $ha^{-1}$, whereas simulated yields from single years and models differed considerably depending on the data aggregation. This implies that large-scale crop yield simulations are robust against climate data aggregation on average. However, they can be systematically biased at higher temporal or spatial resolutions, depending on the model and its parametrization.

**Key words**: Spatial aggregation effects, climate, crop simulation model, input data, scaling, variability, yield simulation, model comparison

# 1) Introduction

Process-based crop models have typically been developed for the field-scale, for which model driving variables (e.g. soil variables) are easily obtained (Van Ittersum et al. 2003, Hansen et al. 2006). However, crop models are increasingly used for large-scale simulations (Chipanshi et al. 1998, Folberth et al. 2012). Scale in the following refers to the spatial extent and resolution of a given grid cell size, ignoring temporal scales (van Bussel et al. 2011a, Weihermüller et al. 2011). Thus, using field-scale crop models with input data at scales other than they were developed for, raises the question how the choice of scale influences the simulation outputs. Changing the spatial resolution by aggregation or disaggregation of data bears the risk of missing the relevant scale of a process or phenomenon, since these are often scale dependent (Meentemeyer 1989). Thus, it is essential to determine the impact of input data (dis-)aggregation on crop model outputs.

Although the relevance of scale (Hansen & Jones 2000, Ewert et al. 2011, Nendel et al., 2013) and spatial data aggregation (Gardner et al. 1982, Cale et al. 1983, Cale & O'Neill 1988, Rastetter et al. 1992, Pierce & Running 1995, Nungesser et al. 1999, Gong et al. 2003, Syphard & Franklin 2004, Lorite IJ et al. 2005, Ershadi et al. 2013) are well-known and data aggregation has been addressed for instance in soil or hydrological process modelling (Heuvelink & Pebesma 1999, Haverkamp et al. 2005, Leopold et al. 2006, Bormann et al. 2009), few studies have characterized the impact in application of crop models with spatially aggregated climate input data on simulated regional yields, hereinafter called aggregation effect. For instance, De Wit et al. (2005) used precipitation and radiation aggregated from 10 km to 50 km resolution as model input to simulate winter wheat and grain maize yields in Germany and France. These yields showed a root mean square error (RMSE) of 0.33 t $ha^{-1}$ ($R^2$ >0.96) and a low bias between results from 10 and 50 km resolutions (estimated from Fig. 4, De Wit et al. 2005). This low bias is in agreement with findings of Folberth et al. (2012) who found small differences in maize yields in the U.S. within resolutions of 7.5 to 45 km resolution. Also Angulo et al. (2013) found small differences in spring barley yields in Finland within resolutions of 10

to 100 km. Similarly, Van Bussel et al. (2011) reported a low bias in simulated winter wheat phenology using aggregated temperatures and sowing dates. Thus, aggregation of climate input data at resolutions of 10 to 100 km could be expected to have low impact on crop model output. This should be considered for instance in large-scale simulations, since data aggregation may destroy physical consistency (Hoffmann & Rath 2012) or combine data beyond the meaningful range of the underlying process. The latter may be taken into account by keeping aggregated grid size below the range estimated via semivariogram models (Brown et al. 1992, Artan et al. 2000). While this previous research indicates the spatial resolution at which crop models may be applied without larger errors in the average simulated crop yields, this research only represented a small sample of crop models and output variables. Furthermore, the systematic behaviour of aggregation effects of models across scales, production situations or crops has not been quantified beyond reporting magnitude and distribution.

We hypothesize that crop models differ in their sensitivity to climate input data aggregation as well as in the fraction of explained variance of the aggregation effects. It is however unknown, whether this affects regional yield estimates from simulations in a systematic way. The objective of this work is therefore to compare the response of regional yields simulated by crop models to climate input data aggregation and to propose a measure for the systematic proportion of the aggregation effects.

# 2) Methods

### 2.1) General procedure and regional focus

The hypotheses given above were tested in the state of North-Rhine Westphalia (NRW), one of the larger a federal states of Germany with a total area of 34,098 $km^2$ (Cologne district council 2013). NRW is characterized by a humid, temperate climate and heterogeneous topography with elevations between 9 m and 843 m above sea level resulting in several agro-ecological zones with different temperature and rainfall regimes. In order to assess spatial aggregation effects, climate input data were aggregated to spatial resolutions varying between 1 km and 100 km (Fig. 1) and used for driving crop models for simulations of winter wheat (*Triticum aestivum* L.) and silage maize (*Zea mays* L.) for the entire state area (Fig. 2). Aggregation effects were estimated by relating model output variables to varying climate inputs. Finally, main determinants of these aggregation effects were identified by analyzing the relative contributions from the climate variables and separately the relative importance of model variables employing Partial Least Squares Regression (PLS).

### 2.2) Climate data processing, aggregation and characterization

**Climate data.** Time series of daily minimum, mean and maximum air temperature (2 m above ground), precipitation, global radiation, wind speed and relative humidity for the period 1982 to 2011 from 280 daily weather stations, as well as an interpolated grid of 1 km resolution of monthly time series were obtained from the German Meteorological Service (DWD). The monthly grids were combined with the daily weather station data as described by Zhao et al. (2014b). Regional climate properties for the different resolutions are given by Table 1 and for 1 km also given by Van Bussel et al. (2014).

**Grids / Aggregation.** Daily climate data in spatial resolution $\alpha$ of 1 $km^2$ were spatially averaged for four different coarser grid sizes of 10, 25, 50 and 100 km (see the Appendix for equations). Coarser grids were technically set-up starting in the north-west corner of the study region.

**Semivariance**. In order to avoid the aggregation of spatially incoherent data, the climate was characterized by semivariograms, which provide information about the extent of spatial dependency (Brown et al. 1993). This so-called range was estimated by fitting Gaussian (precipitation, temperature) and exponential (global radiation) variogram models to the empirical semivariance (Minsasny 2005, Wackernagel 1995, Webster 2001).

**Sub-regions.** Due to incomplete data coverage beyond NRW, data means of grid cells of resolutions larger than 1 km and thus crossing the boundary of the state may be biased, depending on the grid cell size and data coverage (Fig. S3). Therefore, all calculations were validated with the help of sub-regions, being two cells of 100 km resolution and a spatial data coverage of >80 % (C0:R4 and C1:R3, Fig.1) and five cells of 50 km resolution with a corresponding coverage of 100 % (z50, Fig.1).

### 2.3) Crop Simulations

In our study 13 models currently used in addressing different research questions at various scales were selected (appendix Table 5, Suppl. 2). This model ensemble was used to simulate development, growth and yield for the period 1982 to 2011 for each grid cell at each spatial resolutions for winter wheat and silage maize. If not further specified in the following, yield of winter wheat refers to grain yield whereas silage maize yield is aboveground biomass. Simulations (see appendix Table 6 for an overview) were conducted, and both crops were evaluated for i) potential, ii) water-limited and iii) nitrogen-water-limited production situations; correspondingly limited by i) temperature and radiation, ii) precipitation, temperature and radiation, and iii) nitrogen, precipitation, temperature and radiation (Evans and Fisher 1999; Van Ittersum and Rabbinge 1997). Models were set up with a single soil profile (sandy loam; Table S1) which represents a typical deep cropland soil with high water holding capacity and with a common management for any of the grid cells (Appendix Table 7). Models were calibrated at 1 km resolution, using one typical sowing and one typical harvest date per crop as well as the whole region weighted average winter wheat yield and aboveground biomass of silage maize derived from county statistics respectively from 1999 to 2011 and from 2000 to 2008 (Statistische Ämter des Bundes und der Länder 2013). The county statistics are point data and partially based on expert knowledge and are therefore shown only with the purpose to put results into context.

### 2.4) Model intercomparison / Taylor diagrams

Crop model results were compared via Taylor diagrams (Taylor 2001), presenting the correlation coefficient *R*, centered root-mean-square difference (*RMSD*) and standard deviations σ of all grid cells and years as compared to the model ensemble mean (see the appendix for equations). *R* and *RMSD* show correlation and difference, respectively, of each model to the model ensemble mean. *SD* is shown for each model as well as for the ensemble mean. Statistics were calculated from all grid cells of the 1 km resolution and years, thus showing the model agreement in time and space.

### 2.5) Probability density functions (PDFs)

PDFs were obtained by kernel density estimation with a Gaussian kernel (see Hoffmann & Rath 2013 for equations). In order to assure comparability between the PDF of a given crop, the bandwidth was kept constant to 0.1 t $ha^{-1}$ (winter wheat) and 0.3 t $ha^{-1}$ (silage maize).

### 2.6) Analysis of aggregation effects

**Mean regional effects.** Climate data aggregation effects were evaluated in the climate data itself and in model outputs. For this purpose, regional means and spatial variances of daily climate data

and of annual crop model outputs were calculated as absolute differences from coarser resolutions to the 1 km resolution (see appendix 8.2 for equations). Model outputs were thus analyzed only at the resolution of the input data (e.g. to calculate the spatial variance) as well as the mean of the entire region (e.g. to calculate regional yields).

**Fraction of directed effects**. We tested if winter wheat and silage maize yields follow a specific ascending or descending order related to the order of the spatial resolution, for instance if yields at 100 km resolution were larger than yields at 50 km resolution and if yields at 50 km resolution were larger than yields at 25 km resolution and so on. For this purpose, the fraction *P* of simulated yields following monotonously the order or inverse order of the spatial resolutions for each model was calculated (see appendix 8.2 for equations).

**Partial Least Squares Regression (PLS)**. While *P* gives insights into the direction and behavior of aggregation effects, the relevance of factors (e.g. of single climate variables) to aggregation effects remains unclear. However, while the estimated effects and possible causes will be highly auto-correlated, regression algorithms may fail to establish statistical relations between independent (aggregated input dataset) and dependent (crop model output, e.g. yield) data (Luedeling & Gassner 2012). Additionally, the algorithm must handle a large number of independent variables, while avoiding over fitting. In order to evaluate aggregation effects, which are likely driven by numerous complementary as well as contrary processes of crop model and input data interaction, Partial Least Squares Regression (PLS; also known as Projection to Latent Structures Regression) is applied. PLS has been employed for similar purposes related to impacts of climate variation, i.e. on tree phenology (Luedeling & Gassner 2012, Guo et al. 2013). The method takes the dependent variable into account, selecting only the most relevant linear combinations for regression. We thus introduce PLS as a method for quantifying the fraction of variance of aggregation effects explained by climate data or model outputs. PLS was used to assess the relative importance of the climate variables and a limited number of model variables to the aggregation effects.
Dependent variables were changes in mean yields (winter wheat) and final aboveground biomass (silage maize), whereas i) the climate variables during the entire growing period, from sowing to anthesis and from anthesis to maturity, as well as ii) model outputs (winter wheat grain yield, silage maize aboveground biomass, maximum leaf area index, cumulative evapotranspiration, cumulative intercepted photosynthetic active radiation, duration of phenological phases) were used as independent variables. The importance of these variables to the aggregation effect was estimated calculating the variable importance (VIP) from the PLS loadings (Wold 1994, 2001). The highest five predictors with a VIP above 1 were selected.

(Figure 1)

(Figure 2)

# 3) Results

## 3.1) Characterization of the climate input data: Semivariance and aggregation effects

Spatial aggregation of climate time series removed climate extremes of the region, while area means were stable across resolutions (Fig. 3). Consistently, the corresponding spatial variance of the climate variables decreased with increasing spatial resolution. The climate variables daily mean temperature, global radiation and precipitation exhibit a spatial autocorrelation in the range and above the largest aggregation used in this study (Fig. 4, effective range of 93.9, 122.5 and 161.9 km for precipitation, daily mean temperature and global radiation, respectively; $R^2$: 0.998, 0.997, 0.995, respectively). Aggregating the climate data decreases the semivariance, as shown for precipitation (Fig. 4).

## 3.2) Characterization of simulated crop yields

**Crop-specific yields.** The ensemble means of simulated winter wheat and silage maize yields of the region were in the range of 7.6 to 8.7 t $ha^{-1}$ and 15.4 to 17.6 t $ha^{-1}$, respectively depending on the production condition (Table 2, Fig. 5). On average, simulated yields were above the observed yields of 7.2 t $ha^{-1}$ and 14.3 t $ha^{-1}$, respectively.

**Temporal variability of simulated crop yields.** On average, crop models reproduced the year-to-year variability of simulated yields calculated from county statistics (Fig. 5). However, the majority of the models simulated a larger year-to-year variability in yields than observed. Generally, simulated temporal variations of yields were better for winter wheat than for silage maize. Single year yield distributions are shown exemplarily in Fig. S2.

**Influence of the production situation on simulated crop yields.** Yields decreased consistently from potential to water-limited to nitrogen-water-limited production (Table 2, Fig. 5). While on average water-limited winter wheat and silage maize yields were respectively 0.4 $^{-1}$ and 1.2 t $ha^{-1}$ lower than potential yields, they were additionally 0.6 t $ha^{-1}$ and 0.7 t $ha^{-1}$ lower under nitrogen-water-limited conditions as under water-limited conditions (Table 2). Noticeably, while nitrogen availability was thus limiting yields comparably stronger on average than water limitations, the latter were noticeably stronger in 1996, 2010 and 2011, corresponding to 74 % of the decline in these years as compared to potential conditions.

**Spatial variability of simulated crop yields.** The interquartile range of yields across the region and from all three production situations was in the range (silage maize) and partially above the range (winter wheat) of observations (not shown). The corresponding coefficient of variation of crop yield across the region for the mean of years was up to 8.2 % and 12.9 % for winter wheat and silage maize, respectively. Distributions of simulated yields across the region at 1 km resolution consistently showed a negative skew in the mean (Fig. 6).

**Crop model variability in simulating yields.** Simulated yields differed across models and crops, showing a higher agreement among models for winter wheat than for silage maize (Fig. 7). For winter wheat, most models – with the exception of DailyDayCent and LandscapeDNDC – had a standard deviation across all years and grid cells in the range of 1 t $ha^{-1}$, while the standard deviation of the ensemble mean was about 0.5 t $ha^{-1}$. Similar results were found for silage maize, where models showed a larger spread in the standard deviation (approximate range 1.3 to 3.5 t $ha^{-1}$) compared to the ensemble mean (1.0 to 1.4 t $ha^{-1}$). However, models are more dispersed for

potential silage maize aboveground biomass than for potential winter wheat grain yield, with only Simplace<LINTUL5> exhibiting a larger spatio-temporal standard deviation of biomass than other models. Contrastingly, while simulations for winter wheat varied little between production situations, simulations for silage maize showed an increasing standard deviation and *RMSD* between single models and the ensemble mean of model outputs, from potential to water-limited and nitrogen-water-limited production. However, correlation coefficients revealed a range of agreement of 0 to 0.75 and -0.5 to 0.95 for winter wheat and silage maize, respectively, underpinning the larger agreement between single models and model ensemble for wheat than for maize.

### 3.3) General characterization of aggregation effects

**Crop-specific aggregation effects.** Yield and biomass distributions were distinctly different for the two crops and affected by aggregation (Table 2, Fig. 6). On average, aggregation of climate input data led to an increase of winter wheat grain yields (Fig. 8). Mean aggregation effects up to 0.2 t $ha^{-1}$ were found for both winter wheat and silage maize, thus resulting in a lower relative aggregation effect in relation to the yield for silage maize as compared to winter wheat (Table 2). While the mean and maximum likelihood of yield probability density functions (PDFs) of the ensemble results over the region were hardly affected by aggregation, the width was reduced with increasing aggregation (Fig. 6). Contrary to the net aggregation effect of the region, crops differed in their PDF as aggregation led to a mode at higher yields of winter wheat in contrast to the PDF mode of silage maize.

**Temporal variability of yield aggregation effects.** Single-year aggregation effects followed no clear pattern, as positive and negative aggregation effects were simulated. In some years, the aggregation effect consistently increased or decreased with the resolution (see example in Fig. S2). For instance, the annual mean yield of winter wheat and silage maize of the region followed the order of the spatial resolutions *P* in 50.1 % and 34.9 % of all years in the mean of models, respectively.

**Spatial variability of yield aggregation effects.** With coarser spatial resolutions, grid cells located at the region boundary were increasingly less represented by data (Fig. 1, Fig. S1, Fig. S3). As the number of grid cells decreases with coarser spatial resolution, a higher fraction of grid cells extends beyond the boundary of the region. However, analysing sub-regions of 50 and 100 km resolution, revealed similar patterns for mean aggregation effects compared to the entire region of NRW (Fig. 8). Similar to the mean aggregation effects, extremes of aggregation effects of sub-regions were comparable to those of NRW. For silage maize, however, sub-region C0:R4 showed larger aggregation effects under water-limited conditions, resulting from simulations of the AquaCrop 4.0 model.

**Influence of the production situation on yield aggregation effects**. Aggregation effects were similar in the median for all three production situations (Fig. 8), showing a similar pattern across resolutions. Again, larger aggregation effects were found under water-limited conditions for silage maize.

### 3.4) Model-specific interaction with aggregation effects as influenced by treatments

**Aggregation effect - crop model interaction with crop.** While aggregation effects were similar among the crops for the ensemble mean, they differed largely in their extent between models (Fig. 8, Tables 3 and 4). The range of effects was larger on average for silage maize than for winter wheat.

However, some single models showed larger positive aggregation effects for silage maize (e.g. HERMES), comparable aggregation effects for both crops (e.g. MONICA) or lower net aggregation effects for silage maize (e.g. SIMPLACE<LINTUL5>).

**Aggregation effect - crop model interaction with time.** Crop models differed substantially in their sensitivity to climate input data aggregation when single years are considered (Tables 3 and 4). Years with lowest and largest aggregation effects (Fig. S4) differed among crop models with no clear pattern. In addition, models differed in their fraction of yields following clearly the order of the resolutions, *P*, with models ranging from 20.0 % (EPIC) to 66.1 % (HERMES). As compared to 1.7 % probability for a randomly ascending or descending sequence out of 120 permutations, the results indicate systematic processes. Ranking models from high to low systematic effects for winter wheat yields according to $P$ results in the following order: STICS > APSIM(modified) > APSIM > SIMPLACE<LINTUL5> > LandscapeDNDC > HERMES > APSIM-Nwheat > MONICA > COUP > EPIC > MCWLA > DailyDayCent. For silage maize aboveground biomass the order is: HERMES > LandscapeDNDC > AquaCrop4.0 > DailyDayCent > SIMPLACE<LINTUL5> > STICS > APSIM(modified) > MONICA > EPIC > APSIM.

**Aggregation effect - crop model interaction with space**. Differences in the interaction between models and the choice of the sub-region did not show a clear trend (data not shown). However, the proportion of crop models where the aggregation effects followed the order of the resolution was lower in the sub-regions (*P*: 18.3 %, 21.8 % and 26.6 % for C0:R4, C1:R3 and z50, respectively).

**Aggregation effect - crop model interaction with production situation.** Crop models differed in their sensitivity to climate input data aggregation across production situations (Table 3 and 4). While most models showed no clear trend, aggregation effects followed a specific order in the case of few models: increasing limitations led to more positive aggregation effects in HERMES and LandscapeDNDC for both winter wheat and silage maize and for MONICA in the case of silage maize, whereas APSIM and APSIM(modified) where not altered for both winter wheat and silage maize and MCWLA in the case of winter wheat. For the crops simulated, aggregation effects of SIMPLACE<LINTUL5>, DailyDayCent, COUP and Apsim-NWheat decreased with increasing limitation in the production situation whereas STICS, EPIC and AquaCrop did not alter or showed an increasing range.

### 3.5) Systematic effects in crop model output due to climate input data aggregation and variable importance

Using differences in climate variables and their spatial variances as independent variables for PLS-regression led to model-specific sets of variables which are most determinant for single year aggregation effects when wheat or silage maize yield are simulated (Tables 3 and 4). The explained variance by PLS varied between crops, models and production situations. For instance, the explained variance was 60, 91 and 82 % of the single year aggregation effects in the wheat yields simulated by LandscapeDNDC for potential, water-limited and nitrogen-water-limited production, respectively, whereas SIMPLACE<LINTUL5> showed an opposite trend with 80, 61 and 54 %, respectively. Other models showed constantly high (HERMES, 73 to 80 %) or mid (STICS, 59 to 61 %) ranges of explained variance in aggregation effects of simulated wheat yields.

Key variables that statistically explained the variance for aggregation effects of single years were identified by the variable importance for projection (VIP). Concerning winter wheat yield, most models showed the highest VIP for variables related to radiation, followed by variables related to

345 temperature (e.g. HERMES) or again radiation (e.g. STICS, COUP, APSIM). Changes in the aggregation
346 effects of wheat yields of STICS for instance are apparently driven mainly by radiation in the period
347 before anthesis, followed by the temperature before anthesis. Few models indicated precipitation-
348 related terms as most important, mainly for water-limited and nitrogen-water-limited runs
349 (SIMPLACE<LINTUL5>, LandscapeDNDC, EPIC, MONICA). For silage maize, the climate variables with
350 the highest VIP to explain aggregation effects are precipitation- and temperature-related variables,
351 and less related to radiation. Furthermore, while changes in aggregation effects in winter wheat
352 were mainly related directly to changes in climate variables, changes in aggregation effects in silage
353 maize were related mainly to changes in the spatial variance of individual climate variables. No clear
354 trend was obtained for the importance of model outputs (Tables S1 and S2) to explain the effects of
355 climate input aggregation on either wheat or silage maize yields.
356
357 (Figure 3)
358 (Figure 4)
359 (Figure 5)
360 (Figure 6)
361 (Figure 7)
362 (Table 2)
363 (Table 3)
364 (Table 4)
365
366

# 4) Discussion

## 4.1) Data aggregation effects on regional climate

As expected for homogeneous data, changes in the climate variables due to spatial aggregation did not alter the regional mean significantly, but decreased the spatial variance and the semivariance while narrowing extremes. The range estimated from semivariograms was above or in the range of the largest spatial resolution investigated. Thus, the requirement of aggregating data of spatial coherence was met. However, aggregation effects at resolutions coarser than 100 km were not investigated. Thus, without further analysis and depending on the research question, averaging climate variables should probably be restricted to spatial resolutions up to the semivariogram range, which was 94 to 162 km in this study depending on the climate variable. Climate input data aggregation up to 100 km is supported by Van Bussel et al. (2011) and Angulo et al. (2013).

## 4.2) General effects of climate input data aggregation on regional mean simulated yields and biomass

Simulated regional winter wheat and silage maize yields (mean of region and years) were biased when using aggregated climate input data. However, aggregation effects were small compared to the effect of production situations, year-to-year variability or variations across crop models. Aggregation effects in mean yields up to 0.2 t $ha^{-1}$ as compared to 1 km grid cells are in line with findings of Folberth et al. (2012) who found a decrease in mean yields of approximately 0.18 t $ha^{-1}$ of maize after aggregating input data from approximately 7.5 to 45 km. Also Angulo et al. (2013) found biases in the median yield of up to 0.26 t $ha^{-1}$ (LINTUL-SLIM) and 0.21 t $ha^{-1}$ (WOFOST) due to the use of input data with resolutions ranging from 10 to 100 km, while biases from other models were lower (<0.08 t $ha^{-1}$). Thus, aggregation effects on crop yields are on average low in all studies.

Our results show that aggregation hardly influenced the mean climate conditions, but decreased the variance of the climate data. Consequently, an impact on crop model outputs was expected through non-linear functions in the models. The low overall aggregation effect can however be explained as follows. Firstly, the present study is situated in a region with a humid, temperate climate favourable for crop growth and all simulation runs used a typical cropland soil with high water retention capacity. Thus, changes in present climate extremes only slightly influenced the simulated mean regional yield. Secondly, aggregation effects may partially cancel out (Rastetter et al. 1992) at the grid cell level when several climate input variables are aggregated simultaneously and at the regional level when effects of single grid cells cancel out over the region. However aggregation effects have an impact at the grid cell level or on spatial patterns (Zhao et al. 2014a), which was also observed in our study. In conclusion, the low biases in simulating mean crop yields improve the confidence in applying crop models across scales for mean yield estimates of a given region in humid temperate conditions. However, the remaining biases still add to the simulation error and they are likely to increase with climate data variance and absolute level, increasing aridity of climate conditions, model complexity and sensitivity, whenever non-linear effects do not cancel out. This mean aggregation effect may therefore be best observed in stress situations as well as under near-optimum conditions.

Unlike the mean, yield variance and distribution narrowed down in this study with aggregation. This is partially in contrast to Angulo et al. (2013), who found the range of the yield distribution to be only marginally influenced by spatial resolution of climate input variables. However, the Angulo et al. study was based on climate data of a relatively topographically uniform region in south-western Finland with humid climate conditions. Thus, the discrepancies between influences of aggregation

effects for the two regions (NRW and south-western Finland) on yield distributions are explained by the spatial variance of the climate.

### 4.3) Crop model-specific sensitivity to input climate data aggregation

**Aggregation effect interaction of crop model and crop.** Despite an overall low bias in mean simulated yield or biomass over the region due to climate input aggregation, the presented results showed differences in the model response. For silage maize biomass, some models showed lower sensitivity to the climate input aggregation (e.g. MONICA, effect in maize biomass: <1 t $ha^{-1}$), whereas other models showed stronger aggregation effects of up to 1.9 t $ha^{-1}$, (HERMES, APSIM). Although most models showed nearly equally positive and negative aggregation effects, only in a few models the aggregation effect had a dominant direction over all spatial resolutions (e.g. dominant negative effect of input data aggregation in STICS, SIMPLACE<LINTUL5>). It is questionable, whether these responses are the result of the model structure, or whether they result from model parameterizations. Finally, major differences were found between crops regarding the climate variables, which explain most of the variance of the single year aggregation effects (radiation for winter wheat, temperature for silage maize). The variance in the aggregation effects were explained in most models to approx. 60 % and in few occasions up to >80 % by PLS. While the crop model structure does not change largely between simulating winter wheat and silage maize, these changes in the model sensitivity again emphasize the role of model settings for the aggregation effects.

**Aggregation effect interaction of crop model and time.** Analysis of the aggregation effects at single-model- and single-year level (figures S2 and S4a – S4f) did not reveal general trends, but underlined the high variability in model responses to climate input data aggregation. It can thus be concluded that the year-to-year variability of yields masks the aggregation effects on long-term simulated yields.

**Aggregation effect interaction of crop model and production situation.** Models reacted differently, depending on the production situation. While the minimum and maximum aggregation effects of AquaCrop4.0 showed a threefold increase from potential to water-limited production, no trend was found across production situations with DailyDayCent, APSIM_modified and SIMPLACE<LINTUL5>. However, some models (HERMES, LandscapeDNDC, Monica, EPIC) showed a tendency towards larger negative aggregation effects on silage maize biomass when comparing potential and limited production. Interestingly, the variance in aggregation effects explained by PLS decreased with increasing limitations in the production situation, being on average 70.3, 62.8 and 60.9 % for winter wheat and 57.9, 52.1 and 50.3 % for silage maize under potential, water-limited and nitrogen-water-limited conditions. This indicates that increasing model complexity by adding sub-routines to account for additional processes potentially increases the fraction of aggregation effects, which can be regarded as not systematic, i.e. residual variability or noise. Angulo et al. (2013) proposed model-specific fingerprints in the form of yield probability density functions (PDFs) after finding larger discrepancies between models than between aggregation levels.

While no characteristic fingerprints were found for soil input data aggregation (Angulo et al. 2014), the model-specific fingerprints remain to be validated for climate input data aggregation. Since the model fingerprint certainly is modulated by the model structure, it may be co-determined by further factors like model parametrization. Considering the different aggregation effects from similar models (Supplementary 2: Apsim, Apsim_modified and Apsim-NWheat), aggregation effects seem to be partly the result of model parametrization. This does not support the hypothesis of model-specific fingerprints (Angulo et al. 2013). This is similar to the findings by Gardner et al. (1982), who

assumed – after testing hypothetical models of varying structure - that the level of complexity does not alter the aggregation effect noticeably. Consequently, climate data aggregation effects cannot directly be attributed to a given process when using regional, ensemble or other pooled outputs (e.g. mean of years). The processes causing the aggregation effects must thus be assessed at the process level, before generalizing the findings at coarser scales.

**Model-specific drivers for aggregation effects.** PLS-regression was used to identify possible drivers of the aggregation effects on mean simulated yields, which were largely masked by the spatial and temporal variance. However, possible model candidates for further analysis could be identified since: 1) The explained variance by PLS shows which models exhibit systematic aggregation effects explainable by a low number of factors, 2) the selected predictors identified climate variables and model outputs as relevant key drivers. While the variance in aggregation effects for HERMES for instance were explained by >70 % by PLS, this was only approx. 50 % for DailyDayCent. Similar results were obtained with generalized linear models (GLM, data not shown). Thus, models differ not only in their sensitivity, but also in their systematic component of aggregation effects. However, no clear general trend distinguishing crop models in their drivers for aggregation effects was found. While few models were identified as being influenced more by e.g. temperature or radiation, the attribution of these variables to processes in the model remain to be interpreted. For instance, HERMES showed aggregation effects well approximated by the duration of the growing period, which itself largely depends on crop specific parameters (temperature sum, base temperature).
Although the structure of the models is known, a direct attribution of aggregation effects to the model structure (supplement 2) fails due to the high variability of the effects. For instance, aggregation effects on average were larger for i) simple light interception approaches than for detailed approaches, ii) models accounting for vernalization than models not considering vernalization, iii) yield formation based on harvest index than for yield formation based on other approaches (e.g. partitioning during reproductive stages). Hence, for a deeper understanding of aggregation effects and causal processes the analysis needs to be combined with other multivariate methods (e.g. pattern recognition).

### 4.6) Generalization of findings towards the assessment of aggregation effects

Following Pierce et al. (1995) and Rastetter et al. (1992), the aggregation effect should increase with increasing variance of the input data. Larger climate variability as well as different average climate conditions could lead to different aggregation effect distributions especially under growth limiting conditions. While this depends on the data (type, spatial heterogeneity) it is unknown how the aggregation effects from climate data compare to those from other data types. Most of the spatial yield variability in Germany is caused by soil properties and its interactions with climate. In order to focus solely on climate data aggregation effects, soil was not considered as a factor in this study. Aggregation effects from soil properties may therefore differ (Angulo et al. 2014) from the present findings. Pierce et al. (1995) compared the contributions of aggregated input data to the resulting bias in simulated net primary production, which was by 32 % due to spatially averaging climate data (topography 32 %, vegetation and soils 34 %). This remains to be verified for crop models.

## 5) Conclusions

Spatial aggregation of climate input data caused considerable aggregation effects for single models and in single years. Simulated regional yield estimates (average of region and years) were less affected. Differences in simulated mean regional yields across models and/or production situations

505 or in single year yields were larger than the aggregation error. The mean aggregation effects across
506 models and years of up to 0.2 t $ha^{-1}$ (<3 %) contribute to the uncertainty of the estimate of regional
507 yield and biomass. Nevertheless, it has been shown that the effects are systematic. Crop models
508 differ in their sensitivity to aggregated data, showing different means and distributions of
509 aggregation effects, which also depend on the production situation and the crop. Crop models differ
510 in their systematic component of aggregation effects, regardless of the extent of the aggregation
511 effect. Aggregation effects can be attributed to different sources, including climate, climate
512 variability and model structure. Having studied a region in which precipitation rarely limits crop
513 growth, global radiation and temperature were identified as the relevant climate variables, which
514 strongly influence the aggregation effects on winter wheat and silage maize yields, respectively.

515 ## 6) Acknowledgment

516 This study was supported by the BMBF/BMELV project on "Modeling European Agriculture with
517 Climate Change for Food Security (MACSUR)" (grant no. 2812ERA115). The work was partly
518 supported by the Swedish Research Council for Environment, Agricultural Sciences and Spatial
519 Planning and by strategic funding ("Soil-Water-Landscape") from the faculty of Natural Resources
520 and Agricultural Sciences (Swedish University of Agricultural Sciences). We thank ACCAF for financial
521 support and Record team (Eric Casellas) for technical support. We thank the Landesbetrieb
522 Information und Technik Nordrhein-Westfalen for providing regional yield data, and the German
523 Meteorological Service for providing weather data. We thank Professor Per-Erik Jansson (Royal
524 Institute of Technology in Stockholm) for valuable advice linked to the application of the Coup-
525 model. E. Teixeira thanks the Royal Society of New Zealand and the Climate Change Impacts and
526 Implications project for New Zealand (CCII) for financial support to collaborate with the INRES group
527 at the University of Bonn. S. Asseng and D. Cammarano acknowledge support from NOAA-RISA and
528 IFPRI.

529 ## 7) References

530

531 Angulo C, Rötter R, Trnka M, Pirttioja N, Gaiser T, Hlavinka P, Ewert F (2013) Characteristic
532 'fingerprints' of crop model responses data at different spatial resolutions to weather input. Eur J
533 Agron 49: 104-114
534
535 Angulo C, Gaiser T, Rötter RP, Børgesen CD, Hlavinka P, Trnka M, Ewert F (2014) 'Fingerprints' of four
536 crop models as affected by soil input data aggregation. Eur J Agron 61: 35 – 48
537
538 Artan GA, Neale CMU, Tarboton DG (2000) Characteristic length scale of input data in distributed
539 models: implications for modeling grid size. J Hydrol 227: 128-139
540
541 Asseng S, Jamieson PD, Kimball B, Pinter P, Sayre K, Bowden JW, Howden SM (2004) Simulated
542 wheat growth affected by rising temperature, increased water deficit and elevated atmospheric
543 CO2. Field Crop Res 85: 85-102
544
545 Asseng S, Keating BA, Fillery IRP, Gregory PJ, Bowden JW, Turner NC, Palta JA, Abrecht DG (1998)
546 Performance of the APSIM-wheat model in Western Australia. Field Crop Res 57: 163-179
547
548 Bahrenberg G, Giese E, Mevenkamp N, Nipper J (2010) Statistische Methoden in der Geographie.
549 Band 1. Univariate und bivariate Statistik. Borntraeger Science Publishers. Stuttgart. p. 86.
550

551 Bergez JE, Chabrier P, Gary C, Jeuffroy MH, Makowski D, Quesnel G, Ramat E, Raynal H, Rousse N,
552 Wallach D, Debaeke P, Durand P, Duru M, Dury J, Faverdin P, Gascuel-Odoux C, Garcia F (2013) An
553 open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems.
554 Environ Modell Softw 39: 39–49
555
556 Bormann H, Breuer L, Graeff T, Huisman JA, Croke B (2009) Assessing the impact of land use change
557 on hydrology by ensemble modelling (LUCHEM) IV: Model sensitivity to data aggregation and spatial
558 (re-)distribution. Advances in Water Resources 32: 171-192
559
560 Brisson N, Mary B, Ripoche D, Jeuffroy MH, Ruget F, Nicoullaud B, Gate P, Devienne-barret F, Recous
561 S, Tayot X, Plenet D, Cellier P, Machet J, Marc J, Delécolle R (1998) STICS : a generic model for the
562 simulation of crops and their water and nitrogen balances . 1. Theory and parameterization applied
563 to wheat and corn. Agronomie 18: 311–346
564
565 Brisson N, Launay M, Mary B, Beaudoin N (2008) Conceptual basis, formalisations and
566 parameterization of the STICS crop model, Quae. ed.
567
568 Brown DG, Bian L, Walsh SJ (1993) Response of a distributed watershed erosion model to variations
569 in input data aggregation levels. Computers & Geosciences 19: 499-509
570
571 Cale WG, Oneill RV (1988) Aggregation and consistency problems in theoretical-models of
572 exploitative resource competition. Ecol Model 40: 97-109
573
574 Cale WG, Oneill RV, Gardner RH (1983) Aggregation error in non-linear ecological models. Journal of
575 Theoretical Biology, 100(3), 539-550
576
577 Chen C, Wang E, Yu Q (2010) Modeling wheat and maize productivity as affected by climate variation
578 and irrigation supply in North China Plain. Agron J 102: 1037-1049
579
580 Chipanshi AC, Ripley EA, Lawford RG (1998) Large-scale simulation of wheat yields in a semi-arid
581 environment using a crop-growth model. Agr Syst 59: 57-66
582
583 Cologne district council (2013) Nordrhein-Westfalen in Zahlen und Geodaten. URL: www.bezreg-
584 koeln.nrw.de/brk_internet/publikationen/abteilung07/pub_geobasis_nrw.pdf (last fetch:
585 15.12.2014).
586
587 Conrad Y, Fohrer N (2009) Modelling of nitrogen leaching under complex winter wheat and red
588 clover crop rotation on a drained agricultural field. Physics and Chemistry of the Earth 34: 530-540
589
590 Del Grosso SJ, Parton WJ, Mosier AR, Hartman MD, Brenner J, Ojima DS, Schimel DS (2001)
591 Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model.
592 In: Schaffer M (Ed.), Modeling Carbon and Nitrogen Dynamics for Soil Management. CRC Press, Boca
593 Raton, Florida, USA, pp. 303–332.
594
595 Del Grosso SJ, Parton W, Mosier AR, Walsh MK, Ojima D, Thornton PE (2006) DAYCENT national scale
596 simulations of N2O emissions from cropped soils in the USA. J. Environ. Qual. 35: 1451–1460
597

De Wit AJW, Boogaard HL, Van Diepen CA (2005) Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. Agr Forest Meteorol 135: 156-168

Evans LT, Fisher RA (1999) Yield potential: Its definition, measurement, and significance. Crop Science 39: 1544-1551

Ewert F, Van Ittersum MK, Heckelei T, Therond O, Bezlepkina I, Andersen E (2011) Scale changes and model linking methods for integrated assessment of agri-environmental systems

Folberth C, Yang H, Wang X, Abbaspour KC (2012) Impact of input data resolution and extent of harvested areas on crop yield estimates in large-scale agricultural modeling for maize in the USA. Ecol Model 235: 8-18

Gardner RH, Cale WG, Oneill RV (1982) Robust analysis of aggregation error. Ecology 63: 1771-1779

Gong X, Barnston AG, Ward MN (2003) The Effect of Spatial Aggregation on the Skill of Seasonal Precipitation Forecasts. J Climate 16, 3059-3071

Guo L, Dai J, Ranjitkar S, Xu J, Luedeling E (2013) Response of chestnut phenology in China to climate variation and change. Agr Forest Meteorol 180: 164-172

Haas E, Klatt S, Fröhlich A, Werner C, Kiese R, Grote R, Butterbach-Bahl K (2012) LandscapeDNDC: A process model for simulation of biosphere-atmosphere-hydrosphere exchange processes at site and regional scale. Landscape Ecology; DOI: 10.1007/s10980-012-9772-x

Hansen JW, Challinor A, Ines A, Wheeler T, Moron V (2006) Translating climate forecasts into agricultural terms: advances and challenges. Clim Res 33: 27-41

Hansen JW, Jones JW (2000) Scaling-up crop models for climate variability applications. Agr Syst 65: 43-72

Heuvelink GBM, Pebesma EJ (1999) Spatial aggregation and soil process modelling. Geoderma 89: 47-65

Hoffmann H, Rath T (2012) Meteorologically consistent bias correction of climate time series for agricultural models. Theoretical and Applied Climatology 110: 129-141

Hoffmann H, Rath T (2013) Future Bloom and Blossom Frost Risk for Malus domestica Considering Climate Model and Impact Model Uncertainties. Plos One 8 (10). doi: 10.1371/journal.pone.0075033

Holzworth DP, Huth NI, De Voil PG, Zurcher EJ, Herrmann NI, McLean G, Chenu K, Van Oosterom EJ, Snow V, Murphy C, Moore AD, Borwn H, Whish JPM, Verrall S, Fainges J, Bell LW, Peake AS, Poulton PL, Hochman Z, Thorburn PJ, Gaydon DS, Dalgliesh NP, Rodriguez D, Cox H, Chapman S, Doherty A, Teixeira E, Sharp J, Cichota R, Vogeler I (2014) APSIM – Evolution towards a new generation of agricultural systems simulation. Environ Model Softw 62: 327-350

Jansson PE, Karlberg L (2004) Coupled heat and mass transfer model for soil-plant-atmosphere systems, Royal Institute of Technology, Department of Civil and Environmental Engineering, Stockholm, Sweden, 435 pp. , URL:

http://www2.lwr.kth.se/Vara%20Datorprogram/CoupModel/NetHelp/default.htm (last fetch: 15.12.2014)

Jones CA, Kiniry JR (1986) CERES-Maize: A simulation model of maize growth and development. Texas A&M University Press, College Station, Texas, 194

Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, Holzworth DP, Huth NI, Hargreaves G, Meinke H, Hochman Z, Maclean G, Verburg K, Snow V, Dimes JP, Silburn M, Wang E, Brown S, Bristow KL, Asseng S, Chapman SC, McCown RL, Freebairn DM, Smith CJ (2003) An overview of APSIM, a model designed for farming systems simulation. Eur J Agron 18: 267-288

Ershadi A, McCabe MF, Evans JP, Walker JP (2013) Effects of spatial aggregation on the multi-scale estimation of evapotranspiration. Remote Sens Environ 131, 51-62

Haverkamp S, Fohrer N, Frede HG (2005) Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation. Hydrol Process 19, 715-727

Kersebaum KC (2007) Modelling nitrogen dynamics in soil-crop systems with HERMES. Nutrient cycling in agroecosystems 77: 39-52

Kersebaum KC (2011) Special features of the HERMES model and additional procedures for parameterization, calibration, validation, and applications In: L.R. Ahuja and L. Ma (ed.): Advances in Agr Syst Modeling Series 2: 65-94

Kraus D, Weller S, Klatt S, Haas E, Wassmann R, Kiese R, Butterbach-Bahl K (2014) A new LandscapeDNDC biogeochemical module to predict $CH_4$ and $N_2O$ emissions from lowland rice and upland cropping systems. Plant Soil 386: 125-149

Leopold U, Heuvelink GBM, Tiktak A, Finke PA, Schoumans O (2006) Accounting for change support in spatial accuracy assessment of modelled soil mineral phosphorus concentration. Geoderma 130: 368-386

Littleboy M, Silburn D, Freebairn D, Woodruff D, Hammer G, Leslie J (1992) Impact of soil erosion on production in cropping systems. I. Development and validation of a simulation model. Soil Res 30: 757-774

Lorite IJ, Mateos L, Fereres E (2005) Impact of spatial and temporal aggregation of input parameters on the assessment of irrigation scheme performance. J Hydrol 300, 286-299

Luedeling E, Gassner A (2012) Partial Least Squares Regression for analyzing walnut phenology in California. Agr Forest Meteorol 158: 43-52

Meentemeyer V (1989) Geographical perspectives of space, time, and scale. Landscape Ecol 3: 163-173

Minsasny B, McBratney AB (2005) The Matérn function as general model for soil variograms. Geoderma 128: 192-207

Nendel C, Berg M, Kersebaum KC, Mirschel W, Specka X, Wegehenkel M, Wenkel KO, Wieland R (2011) The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. Ecol Model 222: 1614-1625

Nendel C, Wieland R, Mirschel W, Specka X, Guddat C, Kersebaum KC (2013) Simulating regional winter wheat yields using input data of different spatial resolution. Field Crop Res 145: 67-77

Nungesser MK, Joyce LA, McGuire AD (1999) Effects of spatial aggregation on predictions of forest climate change response. Climate Res 11: 109-124

O'Neill RV (1977) Transmutations across hierarchical levels. Conference: Symposium on statistical ecology, College Station, Texas, USA, 25 Jul 1977. OSTI ID: 7083394. Report Number(s): CONF-770730-1, TRN: 77-018130

Parton WJ, Holland EA, Del Grosso SJ, Hartman MD, Martin RE, Mosier AR, Ojima DS, Schimel DS (2001) Generalized model for NOx and N2O emissions from soils. J. Geophys. Res. 106: 17403-17420

Pierce LL, Running SW (1995) The effects of aggregating subgrid land-surface variation on large-scale estimates of net primary production. Landscape Ecology 10: 239-253

Raes D, Steduto P, Hsiao TC, Fereres E (2009) AquaCrop-the FAO crop model to simulate yield response to water: ii. Main algorithms and software description. Agron J 101: 438-447

Rastetter EB, King AW, Cosby BJ, Hornberger GM, Oneill RV, Hobbie JE (1992) Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecol Appl 2: 55-70

Shibu ME, Leffelaar PA, van Keulen H, Aggarwal PK (2010) LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. Eur J Agron 32: 255-271

Statistische Ämter des Bundes und der Länder (2013) Regionaldatenbank Deutschland. https://www.regionalstatistik.de/genesis/online/logon, accessed 08/05/2014

Steduto P, Hsiao TC, Raes D, Fereres E (2009) AquaCrop-The FAO crop model to simulate yield response to water: i. Concepts and underlying principles. Agron J 101: 426-437

Syphard AD, Franklin J (2004) Spatial aggregation effects on the simulation of landscape pattern and ecological processes in southern California plant communities. Ecol Model 180, 21-40

Tao F, Yokozawa M, Zhang Z (2009) Modeling the Impacts of Weather and Climate Variability on Crop Productivity over a Large Area: A New Process-based Model Development, Optimization, and Uncertainties Analysis. Agr Forest Meteorol 149: 831–850

Tao F, Zhao Z (2013) Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. Agr Forest Meteorol 170: 146-166

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research-Atmospheres 106: 7183-7192

Van Bussel LGJ, Ewert F, Leffelaar PA (2011a) Effects of data aggregation on simulations of crop phenology. Agriculture Ecosystems & Environment 142: 75-84

Van Bussel LGJ, Ewert F, Zhao G, Hoffmann H, Wallach D, Constantin J, Raynal H, Klein C, Biernath C, Heinlein F, Tao F, Rötter R, Cammarano D, Asseng S, Elliott J, Glotter M, Nendel C, Kersebaum KC, Specka X, Basso B, Baigorria GA, Romero C (2014) Spatial sampling of weather data for regional crop yield simulations. Agric For Met (submitted)

Van Bussel LGJ, Muller C, van Keulen H, Ewert F, Leffelaar PA (2011b) The effect of temporal aggregation of weather input data on crop growth models' results. Agr Forest Meteorol 151: 607-619

Van Ittersum M, Leffelaar P, Van Keulen H, Kropff M, Bastiaans L, Goudriaan J (2003) On approaches and applications of the Wageningen crop models. Eur J Agron 18: 201-234

Van Ittersum MK, Rabbinge R (1997) Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crop Res 52: 197-208

Vanuytrecht E, Raes D, Steduto P, et al (2014) AquaCrop: FAO'S crop water productivity and yield response model. Environ Model Softw 62:351–360. doi: 10.1016/j.envsoft.2014.08.005

Wackernagel H (1995) Multivariate Geostatistics, Springer.

Wang E, Robertson M, Hammer G, Carberry P, Holzworth D, Meinke H, Chapman S, Hargreaves J, Huth N, McLean G (2002) Development of a generic crop model template in the cropping system model APSIM. Eur J Agron 18: 121-140

Wang E, Zhao Z (2013) Improving APSIM for simulation of temperature response of wheat (APSIM--WheatE). In Alderman et al (ed). Proceedings of the workshop Modeling Wheat Response to High Temperature, CIMMYT, El Batan, Texcoco, Mexico, June 19-21, 2013. p 29.

Webster R, Oliver M (2001) Geostatistics for Environmental Scientists. Wiley & Sons.

Weihermüller L, Huisman JA, Graf A, Herbst M, Vereecken H (2011) Errors in Modeling Carbon Turnover Induced by Temporal Temperature Aggregation. Vadose Zone Journal 10: 195-205.

Williams JR (1995) The EPIC model. In: Singh V.P. (Ed.), Computer models of watershed hydrology. Water resources pubblications, Highland Ranch CO, pp. 909-1000.

Wold S (1994) PLS for Multivariate Linear Modeling. In van de Waterbeemd H (Editor), QSAR: Chemometric Methdos in Molecular Design. Methods and Priciples in Medicinal Chemistry. Weinheim, Verlag-Chemie.

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab 58: 109-130

790 Yeluripati JB, Van Oijen M, Wattenbach M, Neftel A, Ammann A, Parton WJ, Smith P (2009) Bayesian
791 calibration as a tool for initialising the carbon pools of dynamic soil models. Soil Biology &
792 Biochemistry 41: 2579–2583
793
794 Zhao G, Hoffmann H, Van Bussel LGJ, Enders A, Specka X, Sosa C, Yeluripati J, Tao FL Constantin J,
795 Teixera E, Grosz B, Doro L, Zhao Z, Nendel C, Kiese R, Raynal H, Eckersten H, Haas E, Wang E, Kuhnert
796 M, Trombi G, Bindi M, Lewan E, Bach M, Kersebaum KC, Rötter R, Roggero PP, Wallach D, Krauss G,
797 Siebert S, Gaiser T, Ewert F (2014a) Effect of weather data aggregation on regional crop simulation
798 for different crops, production conditions, and response variables. Clim Res (submitted)
799
800 Zhao G, Siebert S, Rezaei EE, Yan C, Ewert F (2014b) Demand for multi-scale weather data for
801 regional crop modelling. Agr For Meteorol (submitted)

# 8) Appendix

8.1) Tables

(Table 5)

(Table 6)

(Table 7)

8.2) Equations

**Grids.** We define a regular grid *R* (http://spatialreference.org/ref/epsg/31467/) with grid points

$$
\begin{aligned}
r &= \{(\lambda_i;\phi_j)|\lambda_{i,j}-\lambda_{i-1,j}=\phi_{i,j}-\phi_{i,j-1}=\alpha\} \qquad (1)\\
\lambda &: \text{gauss-krüger northing value [km]}\\
\phi &: \text{gauss-krüger easting value [km]}\\
i,j &: \text{grid index } [-]\\
\alpha &: \text{resolution [km]}
\end{aligned}
$$

Grids of aggregated data were obtained by taking the spatial mean of data at α = 1 km resolution:

$$
\begin{aligned}
f_\theta^\beta(\mathbf{X}) &= \mathbf{X}' \text{ with} \qquad (2)\\
\left\{\mathbf{X}'|x'_{s,z} = \frac{\sum_{i=\beta(s-1)+1}^{i=\beta s}\sum_{j=\beta(z-1)+1}^{j=\beta z} x_{i,j}}{\beta^2}\right\} &\text{ with}\\
X &: \text{grid to be aggregated } (\alpha{=}1)\\
X' &: \text{aggregated grid}\\
\beta &: \text{resolution after aggregation [km]}\\
i,j &: \text{grid indices at } \alpha = 1 \text{ km}\\
s,z &: \text{grid indices at } \beta
\end{aligned}
$$

Five grids of resolutions of 1, 10, 25, 50 and 100 km with the corresponding number of grid cells 34168, 20, 80, 24 and 9 were constructed. Empty grid cells (data unavailable) were ignored in the calculations. Having time series of daily time steps, $f_\theta$ was applied on each time step. The standard distance of the resulting grids, being a measure for the grid point dispersion (Bahrenberg et al. 2010), is given by Fig. S1. Maps displaying main regional climate variables are given by Zhao et al. (2014a).

**Variable notation.** A simplified notation will be used in the following, indexing variables in their dimensions $\Omega = \{A_1, A_2 ... A_k\}$. In the following, equations are conducted over all elements of a given dimension and are applied to all dimensions indicated. For instance, averages are given by

$$
\mu_{A_1,A_2...A_3}(x) =: \frac{1}{\prod_{k=1}^{k} A_k}\sum_{a_1=1}^{A_1}\sum_{a_2=1}^{A_2}\cdots\sum_{a_k=1}^{A_k} x_{a_1,a_2...a_k} \qquad (3)
$$

where the variable $x$ is averaged over dimensions *A*. In the following, these dimensions were considered

$$
\begin{aligned}
\Omega' &= \{\alpha, c, t, m\} \text{ where} \qquad (4)\\
\Omega' &: \text{dimension, array of dimensions}\\
\alpha &: \text{resolution [km]}\\
c &: \text{grid cell [-]}\\
t &: \text{time step [d],[yr]}\\
m &: \text{crop model [-]}
\end{aligned}
$$

e.g. $Yield_{\alpha,c,t,m}$ refers to a yield of a given resolution, grid cell, year and crop model.

**Analysis of aggregation effects.** Mean aggregation effects were calculated as:

$$
\begin{aligned}
\Delta_{c,t}(x_\alpha) &= \mu_{c,t}(x_{\alpha,c,t}) - \mu_{c,t}(x_{\alpha=1,c,t}) \text{ with} \qquad (4)\\
\Delta_{c,t}(x_\alpha) &: \text{Aggregation effect for resolution } \alpha \text{ and variable } x\\
c &: \text{grid cell [-]}\\
t &: \text{time step [d], [yr]}\\
\alpha &: \text{resolution}
\end{aligned}
$$

Eq. 4 was applied on climate data with daily time steps (t, [d]) as well as on yearly yields (t, [yr]) of each model. Consistently, having the spatial variance

$$
\sigma^2(x_\alpha) = \mu_{c,t}\left(\left\{x_{c,t,\alpha} - \mu_c(x_{c,t,\alpha})\right\}^2\right) \qquad (5)
$$

aggregation effects in the spatial variance were quantified as

$$
\Delta(\sigma^2(x_\alpha)) = \sigma^2(x_\alpha) - \sigma^2(x_{\alpha=1}) \text{ with} \qquad (6)
$$

**Analysis of systematic effects in simulated yields.** The order of mean changes in a given model output due to input data aggregation was analyzed by calculating the fraction of simulated yields following monotonously the order of the resolutions for each model:

$$
\begin{aligned}
P &= \mu(\gamma_t) \cdot 100 \text{ , with}\\
\gamma_t &= \begin{cases} 1 & if \quad Y_{t,\alpha=100} > Y_{t,\alpha=50} > Y_{t,\alpha=25} > Y_{t,\alpha=10} > Y_{t,\alpha=1} \\ 1 & if \quad Y_{t,\alpha=100} < Y_{t,\alpha=50} < Y_{t,\alpha=25} < Y_{t,\alpha=10} < Y_{t,\alpha=1} \\ 0 & else \end{cases}\\
Y_{\alpha,t} &= \mu_c(x_{\alpha,c,t})\\
P &: \text{Percentage of yields following monotonously the order of resolutions } \alpha \text{ [\%]}\\
Y &: \text{Winter wheat grain yield of silage maize aboveground biomass [t ha}^{-1}\text{]}\\
\gamma &: \text{counting variable}
\end{aligned}
$$

**Equations for Taylor diagrams.**

The centered root-mean-square difference (RMSD) was calculated as follows:

$$RMSD_m = \left[\frac{1}{N}\sum_{c,t,m}\left(\left(x_{c,t,m}-\mu_{c,t}(x_{c,t,m})\right)-\left(\mu_m(x_{c,t,m})-\mu_{m,c,t}(x_{c,t,m})\right)\right)^2\right]^{1/2} \quad (8)$$

$c$ : grid cell [-]

$t$ : time step [yr]

$m$ : model

$N$ : 990872 (winter wheat), 1025040 (silage maize)

The standard deviation ⅋ for single models and for the model ensemble mean were calculated as described in eq. 5. The correlation coefficient *R* was calculated as follows:

$$R_m = \frac{\frac{1}{N}\sum_{c,t,m}\left(x_{c,t,m}-\mu_{c,t}(x_{c,t,m})\right)\cdot\left(\mu_m(x_{c,t,m})-\mu_{m,c,t}(x_{c,t,m})\right)}{\mu_{c,t}\left(\left\{x_{c,t,m}-\mu_c(x_{c,t,m})\right\}^2\right)\cdot\mu_{c,t}\left(\left\{\mu_m(x_{c,t,m})-\mu_{c,m}(x_{c,t,m})\right\}^2\right)} \quad (9)$$

845 **Tables**

846 Table 1. Climate of North-Rhine Westphalia (1982-2011, not area weighted). Standard deviations are
847 calculated from all grid cells and years of a corresponding resolution.

| Scale | Annual precipitation sum | | Mean daily temperature | | Mean annual global radiation sum | |
|---|---|---|---|---|---|---|
| [km] | Mean [mm $yr^{-1}$] | σ [mm $yr^{-1}$] | Mean [°C] | σ [°C] | Mean [MJ $m^{-2}$ $yr^{-1}$] | σ [MJ $m^{-2}$ $yr^{-1}$] |
| 1 | 898.9 | 214.0 | 9.7 | 1.2 | 3758.0 | 169.4 |
| 10 | 881.7 | 204.0 | 9.6 | 1.2 | 3757.1 | 166.8 |
| 25 | 873.1 | 191.9 | 9.5 | 1.2 | 3756.6 | 163.8 |
| 50 | 853.8 | 170.8 | 9.4 | 1.1 | 3754.0 | 162.0 |
| 100 | 824.4 | 149.8 | 9.4 | 1.0 | 3765.5 | 160.6 |

848

849

850

851

852 Table 2. Impact of climate input data aggregation on model ensemble mean and spatial variance of
853 simulated yields.

| Production situation* | Scale | Winter wheat grain yield | | Silage maize aboveground biomass | |
|---|---|---|---|---|---|
| [-] | [km] | mean [t $ha^{-1}$] | $\sigma^2$ [t $ha^{-1}$]$^2$ | mean [t $ha^{-1}$] | $\sigma^2$ [t $ha^{-1}$]$^2$ |
| P | 1 | 8.6 | 0.3 | 17.4 | 4.2 |
| P | 10 | 8.6 | 0.3 | 17.4 | 3.8 |
| P | 25 | 8.6 | 0.2 | 17.4 | 3.3 |
| P | 50 | 8.6 | 0.2 | 17.5 | 2.5 |
| P | 100 | 8.7 | 0.1 | 17.6 | 1.5 |
| W | 1 | 8.2 | 0.5 | 16.2 | 4.1 |
| W | 10 | 8.2 | 0.5 | 16.1 | 3.7 |
| W | 25 | 8.3 | 0.4 | 16.2 | 3.2 |
| W | 50 | 8.3 | 0.3 | 16.3 | 2.4 |
| W | 100 | 8.4 | 0.2 | 16.4 | 1.6 |
| N | 1 | 7.6 | 0.4 | 15.5 | 4.0 |
| N | 10 | 7.6 | 0.4 | 15.4 | 3.6 |
| N | 25 | 7.6 | 0.3 | 15.5 | 3.1 |
| N | 50 | 7.6 | 0.2 | 15.6 | 2.3 |
| N | 100 | 7.7 | 0.2 | 15.7 | 1.6 |

855 * P: Potential; W: Water-limited; N: Nitrogen-water-limited

856

Table 3. Climate variables related to aggregation effects with winter wheat grain yield. Nomenclature: **d**: difference; **v**: spatial variance; **P**: precipitation sum, **Tmin/Tmean/Tmax**: minimum, mean, maximum air temperature; **R**: global radiation sum; **GP**: growing period (=sowing to maturity); **SA**: sowing-to-Anthesis; **AM**: anthesis-to-maturity; **L30A/L30M**: period of 30 days before anthesis, maturity. Examples: dRSA, difference in the global radiation sum from sowing to anthesis; dvTminGP, difference in the spatial variance of the daily minimum temperature during the growing period. Variables are sorted from one to five in the order of their variable importance.

| Model | Production* Situation | Single year aggregation effect (difference in grain yield, [t ha$^{-1}$]) minimum | maximum | Explained Variance | Variable 1 | Variable 2 | Variable 3 | Variable 4 | Variable 5 |
|---|---|---|---|---|---|---|---|---|---|
| HERMES | P | -0.04 | 0.37 | 0.80 | dRGP | dTminAM | dRL30A | dTmeanAM | dTmeanGP |
| HERMES | W | -0.14 | 0.57 | 0.73 | dRGP | dvTmeanSA | dvTmeanGP | dTmeanGP | dvTmaxGP |
| HERMES | N | -0.13 | 0.58 | 0.74 | dRGP | dvTmeanSA | dvTmeanGP | dTmeanGP | dvTmaxGP |
| MONICA | P | -0.33 | 0.47 | 0.75 | dRGP | dvPSA | dvTmaxSA | dvTmeanSA | dvPGP |
| MONICA | W | -0.60 | 0.47 | 0.59 | dvPSA | dvPGP | dTmaxSA | dvTmeanSA | dvTmaxSA |
| MONICA | N | -0.60 | 0.47 | 0.59 | dvPSA | dvPGP | dTmaxSA | dvTmeanSA | dvTmaxSA |
| SIMPLACE<L5> | P | -0.06 | 0.37 | 0.80 | dRSA | dRGP | dPL30A | dvTmeanSA | dvTminGP |
| SIMPLACE<L5> | W | -0.77 | 0.30 | 0.61 | dvPGP | dvPSA | dTmeanSA | dvTmeanL30M | dvTminL30M |
| SIMPLACE<L5> | N | -0.73 | 0.37 | 0.54 | dvPGP | dvPSA | dTmeanSA | dvTminL30M | dvTmeanL30M |
| STICS | P | -0.60 | 0.62 | 0.59 | dRSA | dRL30A | dTmaxL30A | dRGP | dTmeanL30A |
| STICS | W | -0.58 | 0.58 | 0.61 | dRSA | dRL30A | dTmaxL30A | dvTminGP | dvTmaxSA |
| STICS | N | -0.56 | 0.62 | 0.60 | dRSA | dRL30A | dTmaxL30A | dTmeanL30A | dvTmaxSA |
| MCWLA | P | -0.72 | 0.73 | 0.64 | dRSA | dRGP | dRL30A | dPGP | dvTmeanGP |
| MCWLA | W | -0.72 | 0.73 | 0.64 | dRSA | dRGP | dRL30A | dPGP | dvTmeanGP |
| DayCent | P | -0.51 | 0.49 | 0.52 | dRL30M | dvRL30M | dPL30M | dTminGP | dTmaxL30M |
| DayCent | W | -0.48 | 0.47 | 0.35 | dPL30M | dvPGP | dTminL30M | dTmaxGP | dTmaxL30M |
| DayCent | N | -1.90 | 0.18 | 0.73 | dRL30M | dRGP | dvTmaxGP | dvTmeanGP | dvRL30M |
| LandscapeDNDC | P | -2.36 | 0.17 | 0.60 | dPL30M | dvPGP | dRL30M | dvPL30M | dvTmaxGP |
| LandscapeDNDC | W | -0.88 | 0.18 | 0.91 | dPL30M | dvPGP | dRL30M | dvTmaxGP | dRGP |
| LandscapeDNDC | N | -0.70 | 0.49 | 0.82 | dvPGP | dPL30M | dRL30M | dvPL30M | dvTmaxGP |
| COUP | P | -0.04 | 0.51 | 0.71 | dRL30A | dRGP | dvTminGP | dRSA | dTmaxL30A |
| COUP | W | -0.48 | 0.30 | 0.45 | dRSA | dRL30A | dRL30M | dvTminGP | dvPL30A |
| APSIM | P | -0.03 | 0.39 | 0.77 | dRGP | dRSA | dvTmeanSA | dvTmaxSA | dTmeanSA |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| APSIM | W | -0.03 | 0.41 | 0.77 | dRGP | dRSA | dvTmeanSA | dTmeanSA | dvTmaxSA |
| APSIM | N | -0.13 | 0.27 | 0.57 | dRSA | dRGP | dRL30A | dvTmaxSA | dvTmeanSA |
| APSIM (modified) | P | -0.08 | 0.84 | 0.79 | dRSA | dRGP | dTminAM | dvTmaxSA | dvTmeanGP |
| APSIM (modified) | W | -0.15 | 0.83 | 0.77 | dRSA | dRGP | dTminAM | dvPSA | dvPGP |
| APSIM (modified) | N | -0.21 | 0.58 | 0.71 | dRSA | dTminAM | dTmaxL30M | dvPGP | dRGP |
| EPIC | P | -0.02 | 0.24 | 0.73 | dRGP | dvTminGP | dvRL30M | dTmeanGP | dRL30M |
| EPIC | W | -0.37 | 0.57 | 0.50 | dRGP | dvTminGP | dvTmaxGP | dTmeanGP | dvTmeanGP |
| EPIC | N | -0.27 | 0.25 | 0.18 | dvPGP | dvTminGP | dPGP | dvRL30M | dTmaxGP |
| APSIM-NWHEAT | P | -0.10 | 0.32 | 0.74 | dTminAM | dPL30A | dRGP | dRSA | dRAM |
| APSIM-NWHEAT | W | -0.50 | 0.32 | 0.61 | dvPSA | dvPGP | dTmaxL30M | dPL30M | dTmeanL30M |

*P: Potential; W: Water-limited, N: Nitrogen-water-limited.

Table 4. Climate variables related to aggregation effects with silage maize. Nomenclature: **d**: difference; **v**: spatial variance; **P**: precipitation sum, **Tmin/Tmean/Tmax**: minimum, mean, maximum air temperature; **R**: global radiation sum; **GP**: growing period (=sowing to maturity); **SA**: sowing-to-Anthesis; **AM**: anthesis-to-maturity; **L30A/L30M**: period of 30 days before anthesis, maturity. Examples: dRSA, difference in the global radiation sum from sowing to anthesis; dvTminGP, difference in the spatial variance of the daily minimum temperature during the growing period. Variables are sorted from one to five in the order of their variable importance.

| Model | Production* Situation | Single year aggregation effect (difference in grain yield, [t ha$^{-1}$]) minimum | maximum | Explained Variance | Variable 1 | Variable 2 | Variable 3 | Variable 4 | Variable 5 |
|---|---|---|---|---|---|---|---|---|---|
| HERMES | P | -0.15 | 1.42 | 0.85 | dvPGP | dTminGP | dvTmaxGP | dPL30M | dTmaxGP |
| HERMES | W | -0.15 | 1.42 | 0.85 | dvPGP | dTminGP | dvTmaxGP | dPL30M | dTmaxGP |
| HERMES | N | -0.06 | 1.97 | 0.83 | dvTmaxGP | dvRL30M | dTminGP | dTminL30M | dvPGP |
| MONICA | P | -0.71 | 0.30 | 0.55 | dvRL30M | dvTmeanL30M | dTmaxGP | dvTminL30M | dRGP |
| MONICA | W | -0.63 | 0.83 | 0.49 | dTmaxGP | dvRL30M | dvPGP | dTmeanGP | dTminGP |
| MONICA | N | -0.63 | 0.83 | 0.49 | dTmaxGP | dvRL30M | dvPGP | dTmeanGP | dTminGP |
| SIMPLACE<L5> | P | -1.24 | 0.68 | 0.49 | dTmaxGP | dTminGP | dvTmaxGP | dvPGP | dPGP |
| SIMPLACE<L5> | W | -1.21 | 0.24 | 0.58 | dTminGP | dvPGP | dvTmaxGP | dvRL30M | dTmaxGP |
| SIMPLACE<L5> | N | -1.20 | 0.42 | 0.46 | dvPGP | dTmaxGP | dTminGP | dPL30M | dvRGP |
| STICS | P | -0.61 | 0.25 | 0.56 | dvPGP | dTminL30M | dRL30M | dTminGP | dTmeanGP |
| STICS | W | -1.58 | 0.40 | 0.30 | dvPGP | dRGP | dPL30M | dvRGP | dvRL30M |
| STICS | N | -1.58 | 0.42 | 0.29 | dvPGP | dRGP | dvRGP | dPL30M | dvRL30M |
| DayCent | P | -0.71 | 1.07 | 0.52 | dvPGP | dTmaxGP | dRL30M | dPGP | dTmeanGP |
| DayCent | W | -0.71 | 1.07 | 0.52 | dvPGP | dTmaxGP | dRL30M | dPGP | dTmeanGP |
| DayCent | N | -0.71 | 1.07 | 0.52 | dvPGP | dTmaxGP | dRL30M | dPGP | dTmeanGP |
| LandscapeDNDC | P | 0.00 | 0.57 | 0.80 | dvPGP | dvPL30M | dPL30M | dRGP | dvTmaxGP |
| LandscapeDNDC | W | 0.00 | 0.60 | 0.79 | dvPGP | dvPL30M | dPL30M | dvTmaxGP | dRGP |
| LandscapeDNDC | N | -0.11 | 0.90 | 0.71 | dvPGP | dRGP | dPL30M | dvTmaxGP | dvPL30M |
| APSIM | P | -0.94 | 1.93 | 0.30 | dPL30M | dTmaxGP | dTmeanL30M | dRGP | dTmaxL30M |
| APSIM | W | -0.81 | 1.93 | 0.29 | dPL30M | dTmaxGP | dTmeanL30M | dRGP | dvPL30M |
| APSIM | N | -0.80 | 1.92 | 0.29 | dPL30M | dTmaxGP | dTmeanL30M | dRGP | dvPL30M |
| APSIM (modified) | P | -1.00 | 1.48 | 0.41 | dTmaxGP | dvPL30M | dTmeanL30M | dTmeanGP | dTminL30M |
| APSIM (modified) | W | -1.00 | 1.48 | 0.40 | dTmaxGP | dvPL30M | dTmeanL30M | dTmeanGP | dvPGP |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| APSIM (modified) | N | -1.00 | 1.48 | 0.39 | dTmaxGP | dvPL30M | dTmeanL30M | dTmeanGP | dTminL30M |
| EPIC | P | -0.30 | 0.80 | 0.57 | dvRL30M | dvTminGP | dvTminL30M | dvTmeanGP | dRGP |
| EPIC | W | -1.04 | 0.98 | 0.55 | dvRGP | dPGP | dPL30M | dvRL30M | dvTmaxGP |
| EPIC | N | -1.04 | 0.98 | 0.55 | dvRGP | dPGP | dPL30M | dvRL30M | dvTmaxGP |
| AquaCrop4.0 | P | -0.40 | 0.55 | 0.74 | dPGP | dvRL30M | dvPL30M | dvRGP | dvTmaxGP |
| AquaCrop4.0 | W | -1.26 | 1.99 | 0.44 | dvTminGP | dvRL30M | dvPGP | dPGP | dRL30M |

*P: Potential; W: Water-limited, N: Nitrogen-water-limited.

Table 5. Crop Models*.

| No. | Model | References |
|---|---|---|
| 1 | APSIM-Nwheat | Asseng et al. 1998, 2004; Keating et al. 2003 |
| 2 | APSIM | Keating et al. 2003, Holzworth et al. 2014 |
| 3 | APSIM, modified | Chen et al. 2010; Keating et al. 2003; Wang et al. 2002 |
| 4 | AquaCrop4.0 | Raes et al. 2009; Steduto et al. 2009; Vanuytrecht et al. 2014 |
| 5 | COUP | Conrad & Fohrer 2009; Jansson & Karlberg 2004 |
| 6 | DailyDayCent | Del Grosso et al. 2001, 2006; Parton et al. 2001; Yeluripati et al. 2009 |
| 7 | EPIC v. 0810 | Williams 1995 |
| 8 | HERMES | Kersebaum 2007, 2011 |
| 9 | LandscapeDNDC | Haas et al. 2012, Kraus et al. 2014 |
| 10 | LINTUL5 | Van Ittersum et al. 2003; Shibu et al. 2010 |
| 11 | MCWLA | Tao et al. 2009, 2013 |
| 12 | MONICA | Nendel et al. 2011 |
| 13 | STICS | Bergez et al. 2013; Brisson et al. 1998, 2008 |

* A more detailed description is given in the supplementary 2.

Table 6. Simulation runs conducted by models.

| Factor | Level |
|---|---|
| Crop | Winter wheat, Silage Maize |
| Production situation | Potential [a], Water-limited [b], Nitrogen-water-limited [c] |
| Resolution [km] | 1, 10, 25, 50, 100 |

[a] growth is limited by temperature and radiation.

[b] growth is limited by precipitation, temperature and radiation

[c] growth is limited by nitrogen, precipitation, temperature and radiation

Table 7. Crop model settings and assumptions.

| Domain | Unit | Winter Wheat | Silage Maize |
|---|---|---|---|
| Sowing date | DOY[a] | 274 | 110 |
| Harvest date | DOY[a] | 213 | 263 |
| Average Yield[b] | t $ha^{-1}$ | 7.2 | 14.3 |
| Max. rooting depth | m | 1.5 | 1.5 |
| Time of ploughing | - | autumn | autumn |
| Planting density | $m^{-2}$ | 400 | 10 |
| Sowing depth | m | 0.04 | 0.06 |
| Initial soil moisture relative to available field capacity[c] | % | 50 | 80 |
| Initial Nmin[d] | kg $ha^{-1}$ | 56 | 56 |
| Nitrogen fertilization | kg $ha^{-1}$ | 130, 52, 26 | 30, 208 |
| Date of fertilization | DOY[a] | 60, 105, 152 | 91, 152 |

[a] Day of the year of a non-leap year. [b] Area weighted average yield derived from county statistics, moisture content: 0 %. [c] Set for each soil layer. [d] Total mineral Nitrogen of the soil profile. Values differ with soil layer.

887 **Figure Captions**

888 Fig. 1. Shape and elevation of the state of North-Rhine Westphalia, Germany. Squares display
889 selected areas used to validate results at state level.

890 Fig. 2. Illustration of climate data resolution ® = 1::100 [km] used as input for crop models.

891 Fig. 3. Area weighted mean and range of key climate variables as well as their spatial variance as
892 affected by data aggregation (Tmin, Tmean, Tmax: daily minimum, mean and maximum temperature
893 respectively). Values are average area means and extremes.

894 Fig. 4. Spatial dependency of selected climate variables in North-Rhine Westphalia. A) Empirical
895 semivariance (dots) and fitted variogram model (solid lines). A Gaussian model was fitted to daily
896 mean temperature (Tmean) and precipitation whereas an exponential model was used for global
897 radiation. B) Empirical semivariance (dots) and fitted Gaussian variogram model (solid lines) for
898 precipitation at varying resolutions α.

899 Fig. 5. Simulated winter wheat grain yield and silage maize aboveground biomass for three
900 production situations at 1 km resolution: Single model and observed 25 to 75 percentile range across
901 the region (shaded areas), model ensemble and observed area weighted mean from county-level
902 statistics (white lines). Areas are plotted with transparency, thus darker areas illustrate coinciding
903 simulation results of several models or coincidence of simulation results with observations.

904 Fig. 6. Probability density functions (pdf) of winter wheat grain yield and silage maize aboveground
905 biomass. Pdfs were estimated from mean grid cell yields and biomass (mean of years) using a
906 Gaussian kernel of bandwidth 0.1 and 0.3 t $ha^{-1}$ for winter wheat and silage maize, respectively.

907 Fig. 7. Taylor diagrams of simulated winter wheat and silage maize yields from respectively 29 and
908 30 years and from 34168 grid cells (α = 1 km), showing: the standard deviation of each model (σ),
909 the correlation between the models (*R)* and the centered root mean square difference to the
910 ensemble mean (RMSD). Denser distributions show smaller diversity among models and vice versa.
911 RMSD and standard deviation are given in t $ha^{-1}$. For each model n = 990,872 and n = 1,025,040 for
912 for winter wheat and silage maize, respectively.

913 Fig. 8. Differences of winter wheat grain yield and silage maize aboveground biomass simulated with
914 aggregated input climate data to yield and biomass simulated with climate time series at 1 km
915 resolution of North-Rhine Westphalia (NRW), two 100 $km^2$ subregions C0:R4 and C1:R3 and one
916 subregion, z50, consisting of five 50 $km^2$ grid cells (see Fig. 1 for regions). The figure displays the min-
917 to-max ranges over the models (shaded and hatched areas) and the ensemble median (thick lines).
918 Values are mean values of 1983 to 2011.

919

920

**Figures**

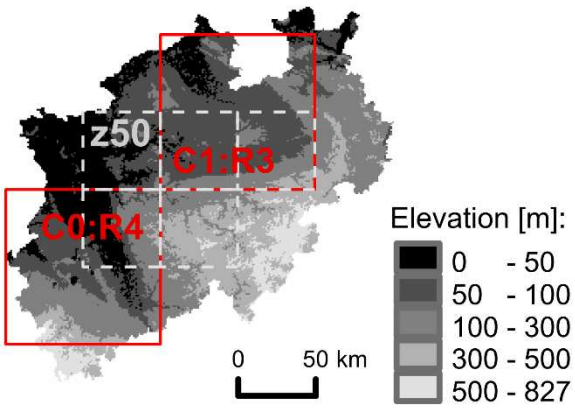



Fig.1

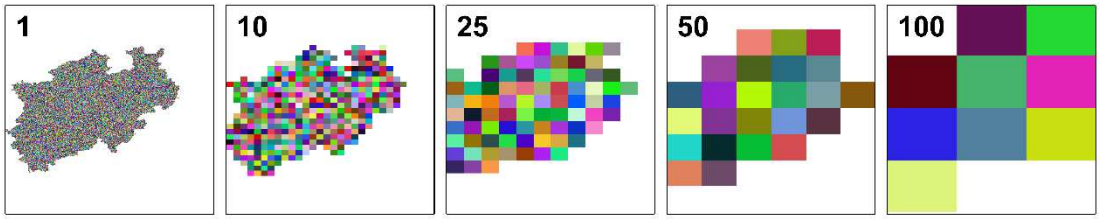
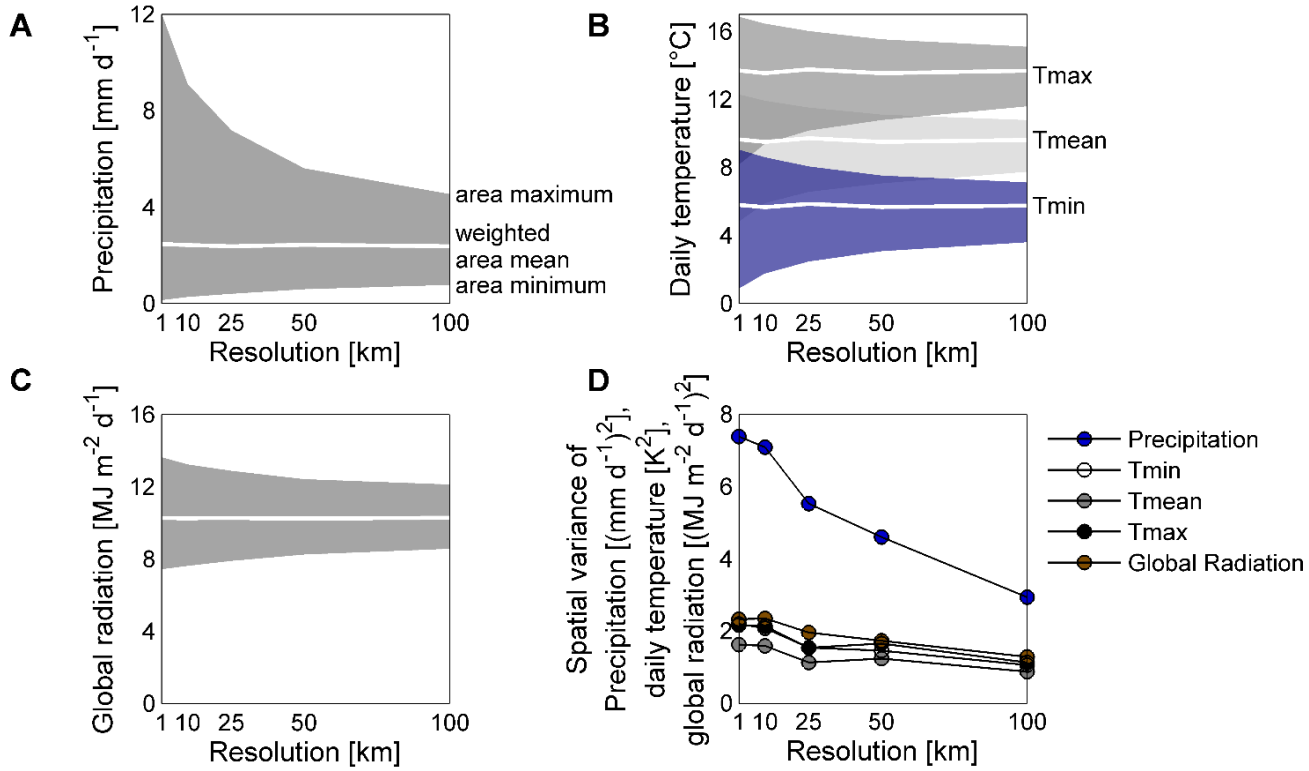



Fig.2




935

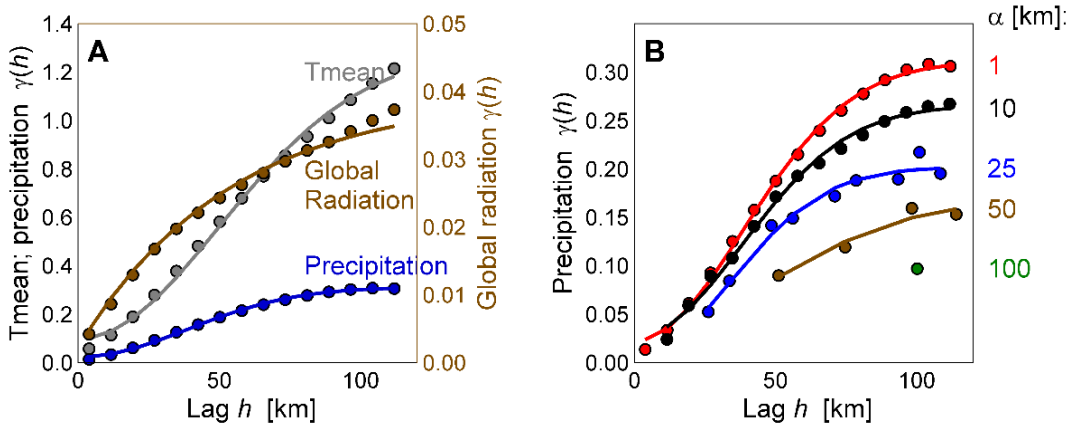
936 Fig. 3

937

938

939

940




941

942 Fig. 4

943

944

945

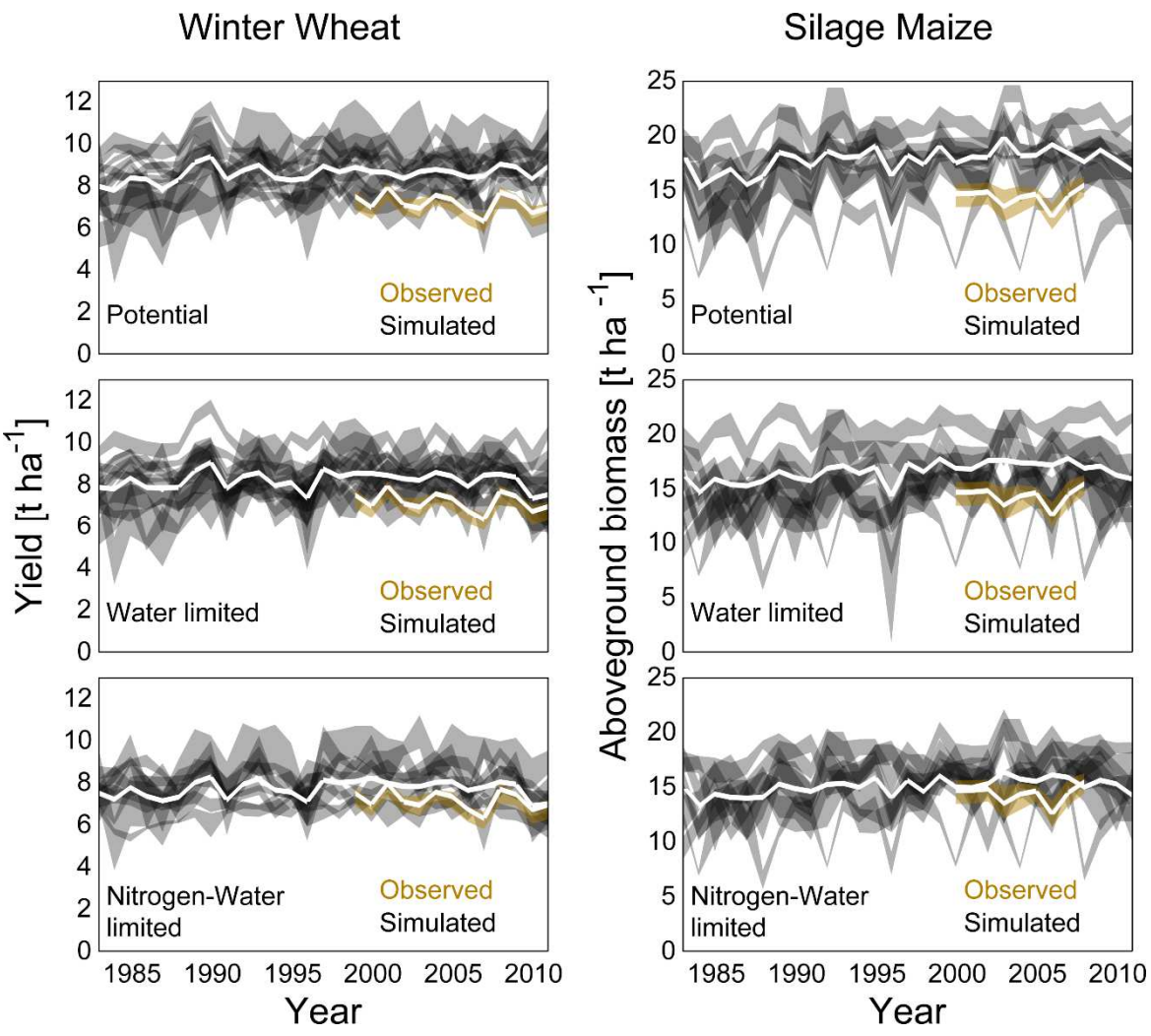

Winter Wheat
Silage Maize
Yield [t ha-1]
Aboveground biomass [t ha-1]
Potential
Water limited
Nitrogen-Water limited
Observed
Simulated
Year
1985
1990
1995
2000
2005
2010

Fig. 5

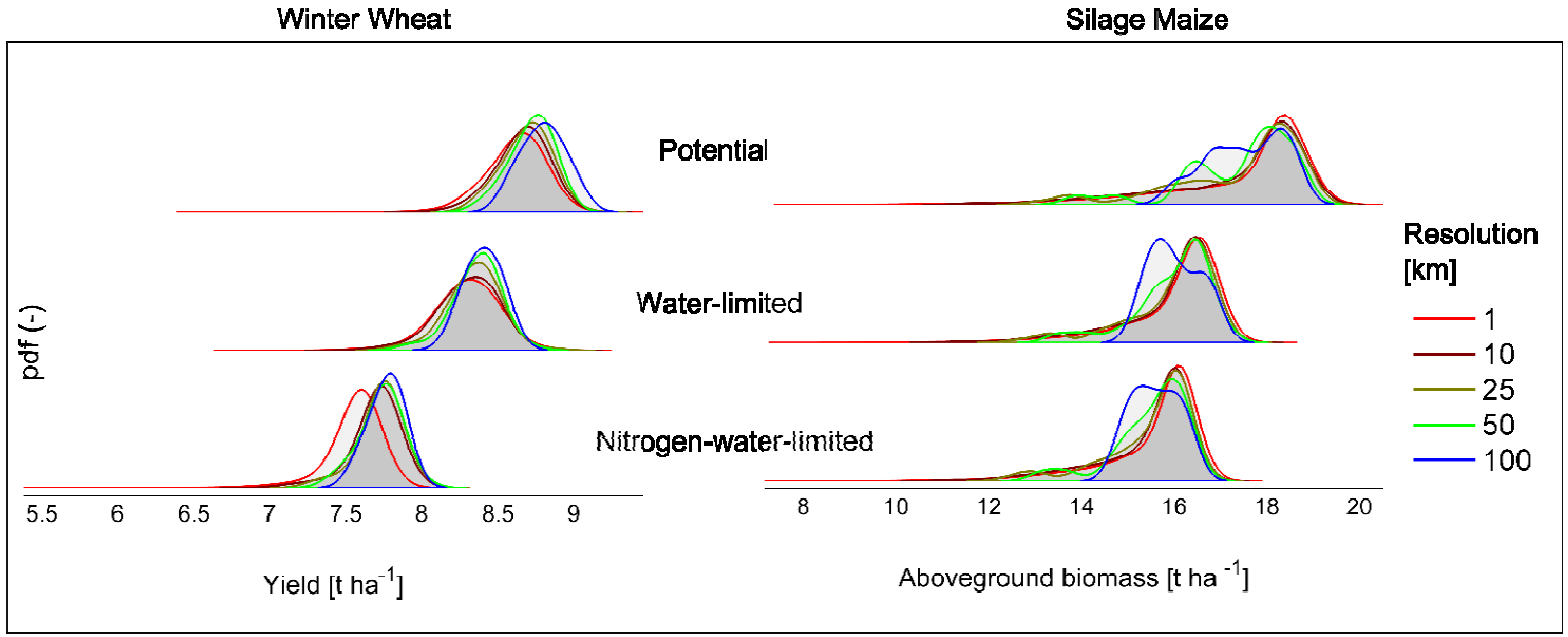
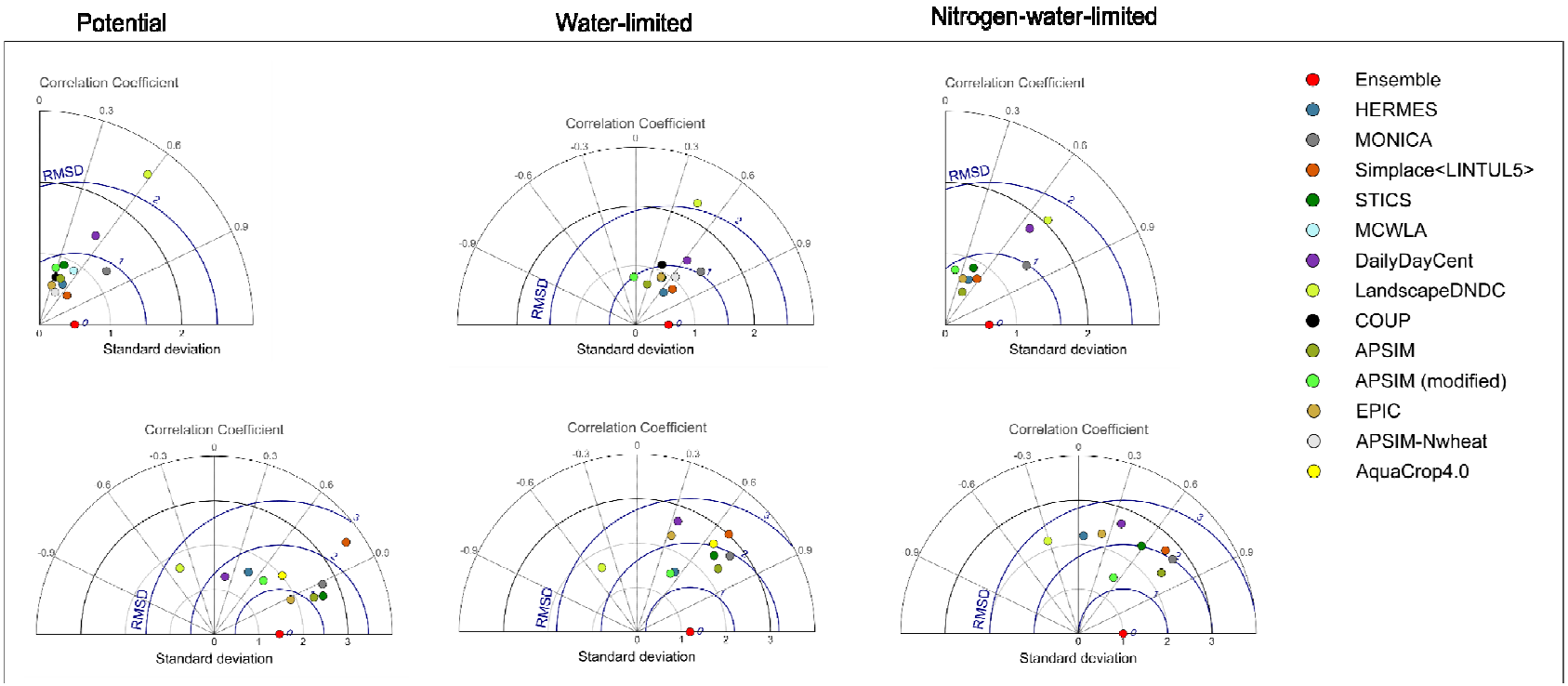

Winter Wheat
Silage Maize
Potential
Water-limited
Nitrogen-water-limited
pdf (-)
5.5
6
6.5
7
7.5
8
8.5
9
Yield [t ha-1]
8
10
12
14
16
18
20
Aboveground biomass [t ha -1]
Resolution [km]
1
10
25
50
100

Fig. 6


Potential
Water-limited
Nitrogen-water-limited
Winter Wheat
Silage Maize
Correlation Coefficient
RMSD
Standard deviation
Ensemble
HERMES
MONICA
Simplace<LINTUL5>
STICS
MCWLA
DailyDayCent
LandscapeDNDC
COUP
APSIM
APSIM (modified)
EPIC
APSIM-Nwheat
AquaCrop4.0

958

959 Fig. 7

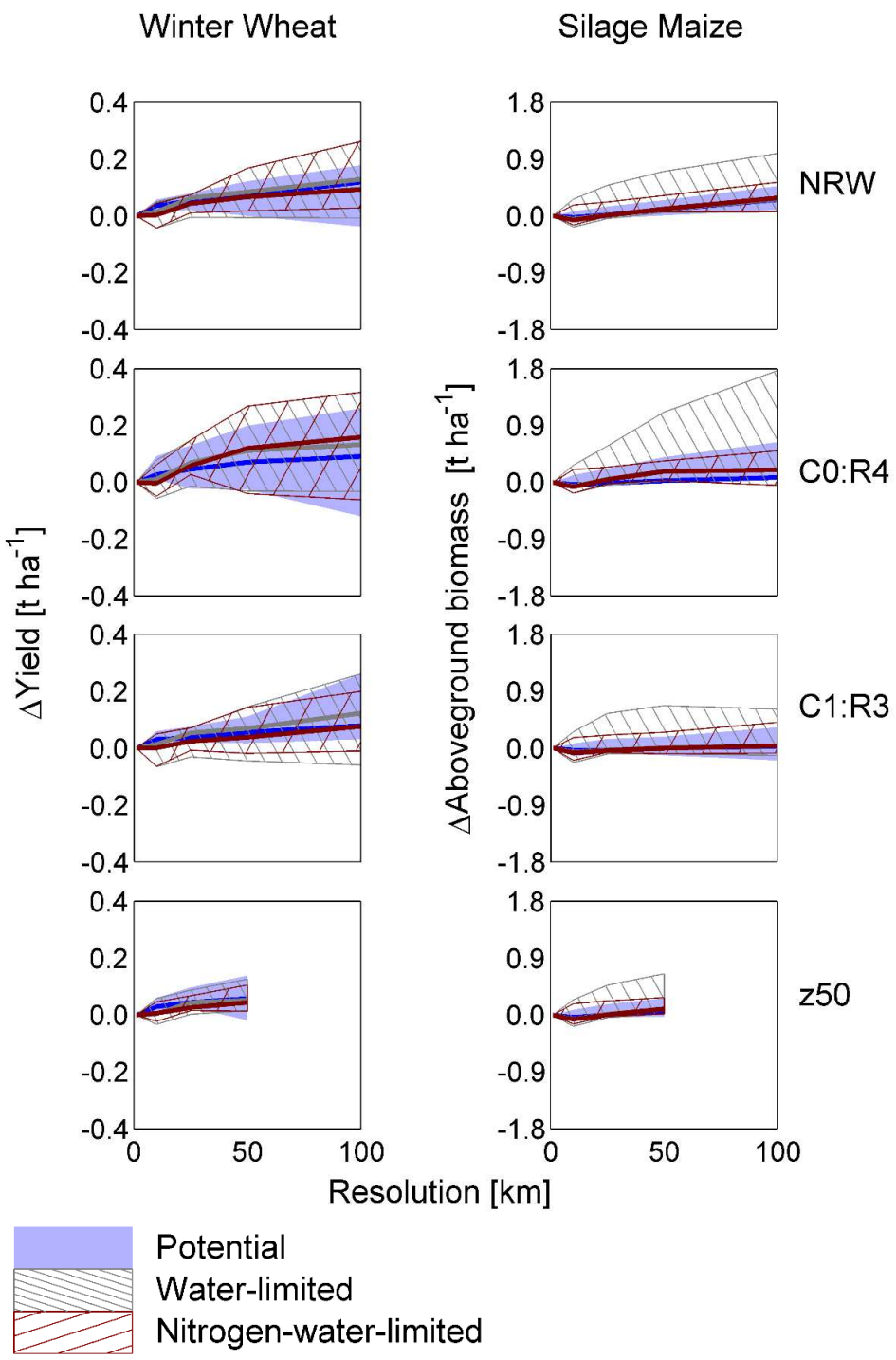

Winter Wheat
Silage Maize
NRW
C0:R4
C1:R3
z50
ΔYield [t ha-1]
ΔAboveground biomass [t ha-1]
Resolution [km]
Potential
Water-limited
Nitrogen-water-limited

Fig. 8